

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 836 779 B1

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention
of the grant of the patent:

03.03.1999 Bulletin 1999/09

(21) Application number: 96923600.9

(22) Date of filing: 02.07.1996

(51) Int Cl. 6: H04L 12/56, H04L 12/46

(86) International application number:
PCT/US96/11200

(87) International publication number:
WO 97/02688 (23.01.1997 Gazette 1997/05)

(54) MECHANISM FOR EFFICIENTLY SYNCHRONIZING INFORMATION OVER A NETWORK

MECHANISMUS ZUR WIRKUNGSVOLLEN SYNCHRONISATION VON INFORMATION IN EINEM
NETZ

MECANISME DE SYNCHRONISATION EFFICACE DE LA CIRCULATION DES INFORMATIONS
DANS UN RESEAU

(84) Designated Contracting States:
DE FR GB IE

(30) Priority: 06.07.1995 US 499029

(43) Date of publication of application:
22.04.1998 Bulletin 1998/17

(73) Proprietor: NOVELL, INC.
Orem, UT 84057-2399 (US)

(72) Inventors:
• PERLMAN, Radia, J.
Acton, MA 01720 (US)
• CASTAGNOLI, Neal, D.
Morgan Hill, CA 95037 (US)

(74) Representative: Hanna, Peter William Derek et al
Tomkins & Co.,
5 Dartmouth Road
Dublin 6 (IE)

(56) References cited:
EP-A- 0 348 331 EP-A- 0 447 725
US-A- 5 265 092

• COMPUTER COMMUNICATIONS REVIEW, vol.
21, no. 1, 1 January 1991, pages 71-78,
XP000175708 ESTRIN D ET AL: "INTER DOMAIN
POLICY ROUTING: OVERVIEW OF
ARCHITECTURE AND PROTOCOLS"

EP 0 836 779 B1

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Description

[0001] This invention relates generally to computer networks and, more particularly, to efficient synchronization of information across a computer network.

BACKGROUND OF THE INVENTION

[0002] A computer network is a geographically distributed collection of interconnected communication links for transporting data between nodes, such as computers. A plurality of computer networks may be further interconnected by intermediate nodes, or *routers*, to extend the effective "size" of the networks. Many types of computer networks are available, with the types ranging from local area networks (LANs) to wide area networks. A LAN, for example, is a limited area network that typically consists of a transmission medium, such as coaxial cable or twisted pair, for interconnecting nodes. These nodes typically communicate by exchanging discrete "packets" of data according to predefined protocols. In this context, a protocol consists of a set of rules defining how the nodes interact with each other.

[0003] In order to reduce design complexity, most networks are organized as a series of hardware and software levels or "layers" within each node. These layers interact to format data for transfer between, e.g., a source node and a destination node communicating over the network. Specifically, predetermined services are performed on the data as it passes through each layer and the layers communicate with each other by means of the predefined protocols. This layered design permits each layer to offer selected services to other layers using a standardized interface that shields those layers from the details of actual implementation of the services.

[0004] In an attempt to standardize network architectures, i.e., the sets of layers and protocols used within a network, a generalized model has been proposed by the International Standards Organization (ISO). The model, called the Open Systems Interconnection (OSI) reference model, is directed to the interconnection of systems that are "open" for communication with other systems. The proposed OSI model has seven layers which are termed, in ascending interfacing order, *the physical, data link, network, transport, session, presentation, and application layers*. These layers are arranged to form a "protocol stack" in each node of the network.

[0005] Fig. 1 illustrates a schematic block diagram of prior art protocol stacks 125 and 175 used to transmit data between a source node 110 and a destination node 150, respectively, of a computer network 100. Each protocol stack is structured according to the OSI seven-layer model; accordingly, each stack comprises a collection of protocols, one per layer. As can be seen, the protocol stacks 125 and 175 are physically connected through a communications channel 180 at the physical layers 124

and 164. For ease of description, the protocol stack 125 will be described.

[0006] Broadly stated, the physical layer 124 transmits a raw data bit stream over a communication channel 180, while the data link layer 122 manipulates the bit stream and transforms it into a datastream that appears free of transmission errors. This latter task is accomplished by dividing the transmitted data into frames and transmitting the frames sequentially, accompanied by error correcting mechanisms for detecting or correcting errors. The network layer 120 routes data packets from the source node to the destination node by selecting one of many alternative paths through the physical network. The transport layer 118 accepts the datastream from the session layer 116, apportions it into smaller units (if necessary), passes the smaller units to the network layer 120 and provides appropriate mechanisms to ensure that all the units arrive correctly at the destination.

[0007] The session layer 116 establishes data transfer "sessions" between software processes on the source and destination nodes, along with management of such sessions in an orderly fashion. That is, a session not only allows ordinary data transport between the nodes, but it also provides enhanced services in some applications. The presentation layer 114 performs frequently-requested functions relating to the presentation of transmitted data, including encoding of data into standard formats, while the application layer 112 contains a variety of protocols that are commonly needed by processes executing on the nodes.

[0008] Data transmission over the network 100 therefore consists of generating data in, e.g., a sending process 104 executing on the source node 110, passing that data to the application layer 112 and down through the layers of the protocol stack 125, where the data are sequentially formatted as a packet for delivery onto the channel 180 as bits. Those packet bits are then transmitted to the protocol stack 175 of the destination node 150, where they are passed up that stack to a receiving process 174. Data flow is schematically illustrated by solid arrows.

[0009] Although actual data transmission occurs vertically through the stacks, each layer is programmed as though such transmission were horizontal. That is, each layer in the source node 100 is programmed to transmit data to its corresponding layer in the destination node 150, as schematically shown by dotted arrows. To achieve this effect, each layer of the protocol stack 125 in the source node 110 typically adds information (in the form of a header field) to the data packet generated by the sending process as the packet descends the stack. At the destination node 150, the various headers are stripped off one-by-one as the packet propagates up the layers of stack 175 until it arrives at the receiving process.

[0010] As noted, a significant function of each layer in the OSI model is to provide services to the other layers.

Two types of services offered by the layers are "connection-oriented" and "connectionless" network services. In a connection-oriented service, the source node establishes a connection with a destination node and, after sending a packet, terminates the connection. The overhead associated with establishing the connection may be unattractive for nodes requiring efficient communication performance. For this case, a fully connectionless service is desirable where each transmitted packet carries the full address of its destination through the network.

[0011] The connectionless network service is generally implemented by a network layer protocol, an aspect of which involves the routing of packets from the source node to the destination node. In particular, this aspect of the network layer concerns the algorithms and protocols used by routers when cooperating to calculate paths through a network topology. A routing algorithm is that portion of the network layer software responsible for determining an output communication link over which an incoming packet should be transmitted; a popular type of network layer routing protocol is a link state routing protocol.

[0012] According to this protocol, each router constructs a link state packet (LSP) comprising information, such as a list of "neighboring" nodes adjacent to the router, sufficient to generate a complete map of the topology of the network. The LSP is then forwarded to all other routers of the network, e.g., a plurality of interconnected LANs. Each of these other routers stores only the most recently received LSP from the forwarding router in its LSP database. Armed with updated maps, these other routers may compute routes to destination nodes. An example of a distributed link state routing protocol is the intermediate system to intermediate system (IS-IS) protocol defined by ISO.

[0013] Since the computed routes are dependent upon the information stored in the LSP databases of the routers, it is essential that these databases are synchronized to ensure their contents are consistent and coherent. A known technique for closely synchronizing LSP databases is to have one node periodically summarize the state of its database. Such technique is for example disclosed in EP-A-0 447 725. According to this technique, which is implemented by the IS-IS protocol, a single router on each LAN of the network is elected a designated router (DR) and the DR periodically transmits a complete sequence numbers packet (CSNP) to all other routers on the LAN. The CSNP consists of identifications of all LSP data items in the database, along with sequence numbers for these items. The routers that receive the CSNP compare it with the contents of their databases to determine whether their information is current.

[0014] For example, if the sequence number of an LSP listed in the CSNP is greater, i.e., more recent, than the sequence number for that LSP stored in the database of a receiving router, that router may request the

more recent information pertaining to the LSP from the DR. On the other hand, if an LSP stored in the database of a receiving router has a sequence number that is greater than the sequence number for that LSP listed in the CSNP, the DR has transmitted "stale" information regarding that LSP. In response to this discovery, the receiving router transmits the more recent information associated with the LSP to the DR, which updates its database. Of course, if the contents of the CSNP are consistent with the contents of the receiving routers' databases, no further action is required.

[0015] In order to characterize an entire LSP database, the CSNP may be very large, thereby requiring apportionment of the CSNP into smaller packet fragments for transmission over the LANs. Each packet fragment characterizes a contiguous portion of the database and each is processed independently by the receiving routers; this enables comparison of the CSNP items with each item of each router's LSP database.

[0016] Therefore, it is among the objects of the present invention to reduce the bandwidth consumed by transmission of database summary information packets over a computer network.

[0017] Another object of the present invention is to minimize computational resources within routers needed to process received database summary information packets.

SUMMARY OF THE INVENTION

[0018] The invention comprises a mechanism for efficiently synchronizing the contents of databases stored on nodes of a computer network to ensure that those contents are consistent. Generally, the mechanism comprises means for generating a database identifier by a node of the computer network and comprises means for distributing to other receiving nodes coupled to the network. The database identifier is uniquely representative of the contents of the distributing node's database and the receiving nodes comprise means for comparing this unique identifier with their own generated database identifiers to determine if the identifiers, and thus their databases, are consistent and synchronized.

[0019] In the illustrative embodiment described herein, the identifier is uniquely representative of a complete sequence numbers packet (CSNP) pertaining to the contents of a link state packet (LSP) database of the distributing node, e.g., a designated router. Specifically, the designated router generates the database identifier from the entire CSNP and periodically broadcasts that identifier, rather than the CSNP itself, to the receiving nodes, i.e., routers, on the network, such as a local area

network (LAN). The database identifier is preferably generated from a cryptographic message digest algorithm configured to transform the contents of the CSNP into a unique, fixed-length digest "signature" whose contents are substantially less than those of the CSNP; accordingly, transmission of the database identifier in lieu of the CSNP optimizes both the use of computational resources within the receiving routers and bandwidth on the LAN.

[0020] Upon receiving the database identifier, the routers process that identifier to determine whether any discrepancies arise and if so, those routers may request copies of the entire CSNP. That is, each receiving router initially calculates an identifier based on the contents of its LSP database and then compares the calculated identifier with the database identifier received from the designated router. A receiving router whose calculated database identifier conforms to the received database identifier need only store that latter identifier of the CSNP. If the calculated identifier is different, the receiving router may request the CSNP to resolve any differences in its database. Significantly, the designated router transmits the actual CSNP only in response to a change in the database or a request from another router.

[0021] In the event a plurality of smaller packet fragments are needed for transmitting the CSNP over the LAN, the designated router preferably computes an identifier for each CSNP fragment. In an alternate embodiment of the invention, a hierarchical arrangement provides a single, high-level database identifier for the entire CSNP and a plurality of low-level database identifiers for these individual CSNP fragments. Here, the high-level identifier is periodically broadcast by the designated router and if a discrepancy is found at a particular router, that router may request a list of the low-level identifiers in order to isolate the discrepancy in the database.

[0022] In a related alternate embodiment, the hierarchical arrangement is modified to provide a two-stage operation arrangement at the receiving routers. Specifically, the high-level and low-level identifiers are bundled within a "hello" message that is periodically broadcast by the designated router to the receiving routers. According to the first operation stage, each receiving router calculates an identifier based on the entirety of its database, compares that identifier with the received high-level identifier and, if they match, ignores the remainder of the message. On the other hand, if the identifiers are dissimilar, the receiving router proceeds to the second stage, which specifies computing identifiers for particular fragments of the database. These latter identifiers are thereafter compared with the appropriate low-level identifiers to identify the inconsistent database fragments.

[0023] Advantageously, the inventive embodiments described above do not require extensive use of computational resources in the receiving routers unless there are inconsistencies in the databases. In other

words, the invention conserves processing resources by potentially eliminating the need to labor through identifier calculations and comparisons for each database fragment. Moreover, these approaches are flexible in that the number of hierarchical layers, database fragments and low-level identifiers are selectable.

BRIEF DESCRIPTION OF THE DRAWINGS

[0024] The above and further advantages of the invention may be better understood by referring to the following description in conjunction with the accompanying drawings, in which like references indicate similar elements, and in which:

15 Fig. 1 is a schematic block diagram of prior art protocol stacks used to transmit data between nodes of a computer network;

20 Fig. 2 is a block diagram of a network system including a collection of computer networks connected to a plurality of nodes;

25 Fig. 3 is a schematic diagram of a conventional link state packet (LSP) used in accordance with a network layer routing protocol;

30 Figs. 4A-4B are schematic diagrams of complete sequence numbers packets used in accordance with a network layer protocol;

35 Fig. 5 is a schematic diagram illustrating an illustrative embodiment of a message containing a novel database identifier mechanism according to the present invention;

40 Figs. 6A-6B are schematic diagrams of alternate embodiments of messages containing high-level and low-level database identifiers in accordance with the invention; and

45 Fig. 7 is a schematic diagram of yet another alternate embodiment of a message containing various level database identifiers in accordance with the invention.

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENT

[0025] Fig. 2 is a block diagram of a network system comprising a collection of computer networks connected to a plurality of nodes. The nodes are typically general-purpose computers comprising source nodes S1-S6, destination node D and intermediate nodes R1-R6. Each node typically comprises a central processing unit (CPU) 202, a memory unit 204 and at least one network adapter 206 interconnected by a system bus 210. The memory unit 204 may comprise storage locations typically composed of random access memory (RAM) devices, which are addressable by the CPU 202 and network adapter 206. An operating system, portions of which are typically resident in memory and executed by CPU, functionally organizes the node by, *inter alia*, invoking network operations in support of

processes executing in the CPU.

[0026] The computer networks included within system 200 are local area networks (LANs) 1-2 interconnected by intermediate node R4, which is preferably a router. Communication among the nodes coupled to the LANs is typically effected by exchanging discrete data "packets" among the nodes. Router R4 facilitates the flow of these data packets throughout the system by routing the packets to the proper receiving nodes.

[0027] In general, when a source node transmits a packet over LAN 1, the packet is sent to all nodes on that LAN. If the intended recipient of the packet is connected to LAN 2, the packet is routed through router R4 onto LAN 2. Typically, the packet contains two destination addresses: the address of the final destination node and the address of the next node along the route. The final destination address remains constant as the packet traverses the networks, while the next destination address changes as the packet moves from node to node along the route through the networks.

[0028] Specifically, when source node S1 sends a packet to destination node D, i.e., the final destination address, the packet is transmitted onto LAN 1 with a next destination address specifying the address of router R4. Address information embedded in the packet, which is processed by the higher-layer software of the protocol stack 250, identifies the final destination of the packet as node D. Based on this information, R4 determines that the next node is the final destination node D and transmits the packet over LAN 2 to node D.

[0029] A key function of a router is determining the next node to which the packet is sent; this routing function is preferably performed by network layer 260 of a protocol stack 250 within each node. This aspect of the network layer concerns the algorithms and protocols used by routers when cooperating to calculate paths through a network topology. The routers typically execute routing algorithms to decide over which communication links incoming packets should be transmitted; a type of network layer routing protocol commonly employed by routers is a link state routing protocol.

[0030] According to this protocol, each router constructs a link state packet (LSP) containing information needed to generate a complete map of the topology of the network. Fig. 3 depicts a schematic diagram of the LSP 300 comprising, *inter alia*, a source field 302 that indicates the particular source node generating the LSP, a sequence number field 304 for storing the sequence number of the LSP and a neighbors field 306 containing a list of "neighbors", i.e., nodes adjacent to the source node. The sequence number is preferably a monotonically increasing value that functions as a counter to uniquely identify the LSP.

[0031] Each router forwards its LSP 300 to all other routers coupled to the network and each of the other routers stores only the most recently received LSP in a LSP database organized in each of their memories 204 (Fig. 2). Armed with an updated set of LSPs, each router

may execute predetermined algorithms to compute routes to destination nodes. An example of a distributed link state routing protocol is the intermediate system to intermediate system (IS-IS) protocol.

5 [0032] Since the computed routes are dependent upon the information stored in the LSP databases of the routers, it is essential that these databases are synchronized to ensure their contents are consistent. A known technique for closely synchronizing LSP databases involves designating a single router on the network as a designated router that periodically transmits a complete sequence numbers packet (CSNP) to all other routers on the network. Fig. 4A is a schematic diagram of a CSNP 400 comprising a list of identifications of LSP data items 410 in the designated router's database, along with sequence numbers 420 for these items.

10 [0033] The routers that receive the CSNP 400 compare it with the contents of their databases to determine whether their information is current. That is, the routers compare each sequence number of the items listed in the CSNP with the sequence number of corresponding data items of their databases to determine if they are equal. If they are not, the greater, i.e., more recent, data item as indicated by the sequence number is provided

15 to the router having the less recent item.

[0034] In order to characterize an entire LSP database, the CSNP may be very large, thereby requiring apportionment of the CSNP into smaller packet fragments for transmission over the network. Fig. 4B depicts schematic diagrams of CSNP fragments 450a-c, each of which comprises a list of sequence numbers tied to an address range of LSP data items. Each packet fragment 450a-c preferably characterizes a contiguous portion of the database, e.g., addresses xx-zz, and each is processed independently by the receiving routers. However, as noted, transmission of these additional smaller packet fragments 450a-c over the network consumes significant bandwidth, while processing of the additional individual packets consumes substantial amounts of computational resources in the routers.

20 [0035] In accordance with the invention, a mechanism is provided for efficiently synchronizing the contents of databases stored on nodes of a computer network to ensure that those contents are consistent. Specifically, 25 the mechanism comprises a database identifier generated by a node of the computer network and distributed to other nodes coupled to the network. According to the invention, the database identifier is uniquely representative of the contents of the node's database and the other nodes compare this unique identifier with their own generated database identifiers to determine if the identifiers, and thus their databases, are consistent.

[0036] In the illustrative embodiment, the database identifier is uniquely representative of the CSNP. Referring also to Fig. 2, the node (e.g., a designated router which, for purposes of description, is router R4) generates the database identifier from the entire CSNP 400 and periodically broadcasts that identifier, rather than

the CSNP itself, to other nodes (e.g., routers) R1-R3 and R5-R6 coupled to LANs 1-2 by way of e.g., a "hello" message. Fig. 5 is a schematic diagram of a hello message packet 500 containing, *inter alia*, information such as the novel database identifier 510 of the invention and the source identification (originator) 502 of the message, i.e., the designated router R4.

[0037] Preferably, the database identifier 510 is generated from a cryptographic message digest algorithm executed by the CPU 202 and configured to transform the CSNP into a unique, fixed-length digest "signature" whose contents are substantially less than those of the CSNP. It should be noted, however, that the identifier 510 may be generated by other techniques, such as cyclic redundancy checking or sequence number generation by the CPU. The underlying requirement of such techniques is that they must be capable of producing unique values with high probability.

[0038] In the illustrative embodiment, the contents of the database identifier field may comprise a 64 or 128 bit length of the message, although any concise signature of relatively modest fixed length would suffice. A significant aspect of the invention is that the routers need only examine and compare the contents of these fixed length fields to ascertain the coherency of their databases. Accordingly, transmission of the database identifier in lieu of the CSNP optimizes both the use of computational resources within the other routers and bandwidth on the network.

[0039] Upon receiving the database identifier, the routers R1-R3 and R5-R6 process that identifier to determine whether any discrepancies arise and if so, those routers may request copies of the entire CSNP from the designated router R4. That is, each receiving router initially calculates an identifier based on the contents of its LSP database and then compares the calculated identifier with the database identifier received from the designated router. Of course, the routers R1-R3 and R5-R6 calculate their identifiers according to the same algorithm or technique used by router R4.

[0040] A receiving router whose calculated database identifier conforms to the received database identifier need only store that latter identifier of the CSNP. If the calculated identifier is different, the receiving router may request the entire CSNP from the designated router R4 to resolve any differences in its database. Significantly, R4 transmits the actual CSNP only in response to a change in its database or in response to a request from another router.

[0041] In the event a plurality of smaller packet fragments 450a-c are needed for transmitting the CSNP 400 over the LANs 1-2, the designated router preferably computes an identifier for each CSNP fragment. In an alternate embodiment of the invention, a hierarchical arrangement provides a single, high-level database identifier for the entire CSNP and a plurality of low-level database identifiers for these individual CSNP fragments. Referring to Figs. 6A and 6B, the high-level identifier 610

is contained within a high-level hello message 600 that is periodically broadcast by the designated router R4. If a discrepancy between identifiers is discovered by a router, that router may request a particular low-level identifier 625a-c stored in low-level messages 620a-c, respectively; each identifier 625a-c corresponds to an appropriate CSNP fragment 450a-c.

[0042] In a related alternate embodiment, the hierarchical arrangement is further modified to provide a two-stage operation arrangement at the receiving routers R1-R3 and R5-R6. Specifically, the high-level and low-level identifiers are bundled within the same hello message that is periodically broadcast by the designated router R4 to the other routers. Fig. 7 is a schematic diagram of a hello message 700 containing the various level identifiers, such as high-level identifier 710 and low-level identifiers 725a-c.

[0043] According to the first operation stage of the arrangement, each receiving router calculates an identifier based on the entirety of its database, compares that identifier with the received high-level identifier 710 and, if they match, ignores the remainder of the message 700. On the other hand, if the identifiers are dissimilar, the receiving router proceeds to the second stage, which specifies computations of identifiers for particular fragments of the database. These latter identifiers are thereafter compared with the appropriate low-level identifiers 725a-c to identify the inconsistent database fragments.

[0044] One advantage of the invention is that extensive use of computational resources in the receiving routers is not required unless there are inconsistencies in the databases. In other words, the invention conserves processing resources by potentially eliminating the need to labor through identifier calculations and comparisons for each database fragment. In addition, the invention is flexible in that the number of hierarchical layers, database fragments and low-level identifiers are selectable.

[0045] While there has been shown and described illustrative embodiments for implementing a mechanism that efficiently synchronizes the LSP databases of routers coupled to a LAN, it is to be understood that various other adaptations and modifications may be made within the spirit and scope of the invention. For example, the mechanism described herein may be used in any type of distributed system requiring efficient synchronization of the contents of databases stored on nodes of a computer network. In the case of such distributed systems, a designated node of the computer network generates the database identifier and distributes that identifier to other nodes coupled to the network. According to the invention, the database identifier is uniquely representative of the contents of the designated node's database and the receiving nodes generate their own database identifiers from the contents of their databases so that they may compare their identifiers with that of the designated node to ensure that the contents of their data-

bases are consistent.

[0046] The foregoing description has been directed to specific embodiments of this invention. It will be apparent, however, that other variations and modifications may be made to the described embodiments, with the attainment of some or all of their advantages. Therefore, it is the object of the appended claims to cover all such variations and modifications as come within the true scope of the invention.

Claims

1. A mechanism for efficiently synchronizing the contents of databases stored on nodes of a computer network (200) characterized in that the mechanism comprises means for generating a database identifier (510) by each node of the computer network from the contents of a database associated with each node, the database identifier uniquely representative of the contents of the database associated with each node, whereby a node (S1 - S6; R1 - R6) of the network comprises means for distributing its database identifier to a receiving node (D) of the network and the receiving node comprises means for comparing its database identifier with the distributed database identifier to determine if the identifiers, and thus the databases, are synchronized.
2. The mechanism of Claim 1, wherein the database identifier is uniquely representative of a complete sequence numbers packet (CSNP) (400) pertaining to the contents of a link state packet (LSP) (300) database of the distributing node (S1 - S6; R1 - R6).
3. The mechanism of Claim 2, wherein the nodes are routers (R1 - R6) and wherein the distributing node is a designated router (R4).
4. The mechanism of Claim 1, wherein the database identifier (510) is generated by a cryptographic message digest algorithm configured to transform the contents of the database into a unique, fixed-length digest, by cyclic redundancy checking, or by sequence number generation.
5. An arrangement for efficiently synchronizing the contents of databases stored on routers (R1 - R6) of a computer network (200) characterized in that the arrangement comprises:

a designated router (R4) coupled to the network and configured to generate a first database identifier from the contents of a first database associated with the designated router, the first database identifier uniquely representative of the contents of the first database; and at least one other router of the network config-

5. ured to generate a second database identifier from the contents of a second database associated with the other router, the second database identifier uniquely representative of the contents of the second database, wherein the designated router periodically distributes the first database identifier to the other router which compares the first database identifier with the second database identifier to determine whether the database identifiers, and thus the first and second databases, are consistent.
10. 6. The arrangement of Claim 5, wherein the database identifiers are uniquely representative of complete sequence numbers packets (CSNP) (400) pertaining to the contents of link state packet (LSP) (300) databases and wherein the designated router generates the first database identifier from the CSNP and periodically broadcasts the first database identifier, rather than the CSNP itself, to the other router.
15. 7. A method for efficiently synchronizing the contents of databases stored on nodes (S1 - S6; R1 - R6) of a computer network (200) characterized in that the method comprises:
20. generating a first database identifier from the contents of a first database associated with a designated node coupled to the network, the first database identifier uniquely representative of the contents of the first database; calculating second database identifiers from the contents of second databases associated with a plurality of other routers coupled to the network, the second database identifiers uniquely representative of the contents of the second databases; periodically broadcasting the first database identifier to the other routers over the network; and
25. processing the first database identifier at the other routers to determine whether there are any discrepancies between the first and second databases.
30. 8. The method of Claim 7, wherein the step of generating comprises the step of generating the first database identifier that is uniquely representative of a complete sequence numbers packet (CSNP) pertaining to the contents of a link state packet (LSP) database associated with the designated router.
35. 9. The method of Claim 8, wherein the step of processing comprises the step of comparing the second database identifier with the first database identifier, and if the second database identifier matches the first database identifier, storing the first database
- 40.
- 45.
- 50.
- 55.

identifier at the other routers, or if the second database identifier does not match the first database identifier, requesting the CSNP from the designated router.

5

10. The method of Claim 8, further comprising the steps of:

generating third database identifiers (625a - c), each of which is uniquely representative of a fragment (450a - c) of the CSNP pertaining to the contents of a LSP database associated with the designated router; and

10

calculating fourth database identifiers, each of which is uniquely representative of a fragment of the CSNP pertaining to the contents of a LSP database associated with each other router.

15. The method of Claim 10, wherein the step of processing comprises the step of comparing the second database identifier with the first database identifier, and if the second database identifier matches the first database identifier, storing the first database identifier at the other routers, or if the second database identifier does not match the first database identifier, comparing a particular third database identifier with a corresponding fourth database identifier to resolve any discrepancies.

20

25

30

Patentansprüche

1. Mechanismus zur wirksamen Synchronisierung der Inhalte von Datenbanken, die auf Knoten eines Computernetzwerks (200) gespeichert sind, dadurch gekennzeichnet, daß der Mechanismus:

35

eine Vorrichtung zur Erzeugung eines Datenbank-Identifizierers (510) aufweist, der von jedem Knoten des Computernetzwerks aus dem Inhalt einer Datenbank erzeugt wird, welche jedem Knoten zugeordnet ist, wobei der Datenbank-Identifizierer eindeutig den Inhalt der Datenbank identifiziert, welche jedem Knoten zugeordnet ist, wodurch ein Knoten (S1-S6; R1-R6) des Netzwerks eine Vorrichtung zum Verteilen seines Datenbank-Identifizierers zu einem Empfangsknoten (E) des Netzwerks aufweist, und der Empfangsknoten eine Vorrichtung zum Vergleichen seines Datenbank-Identifizierers mit dem verteilten Datenbank-Identifizierer aufweist, um festzustellen, ob die Identifizierer und daher die Datenbanken synchronisiert sind.

40

45

50

55

3. Mechanismus nach Anspruch 2, bei welchem die Knoten Leitwegvorrichtungen (R1-R6) sind, und der verteilende Knoten eine festgelegte Leitwegvorrichtung (R4) ist.

4. Mechanismus nach Anspruch 1, bei welchem der Datenbank-Identifizierer (510) von einem Verarbeitungsalgorithmus für eine kryptographische Nachricht erzeugt wird, der so ausgebildet ist, daß er den Inhalt der Datenbank in eine eindeutige Verarbeitung mit fester Länge umwandelt, durch zyklische Redundanzüberprüfung, oder durch Sequenznummererzeugung.

5. Anordnung zum wirksamen Synchronisieren der Inhalte von Datenbanken, die auf Leitwegvorrichtungen (R1-R6) eines Computernetzwerks (200) gespeichert sind, dadurch gekennzeichnet, daß die Anordnung aufweist:

eine festgelegte Leitwegvorrichtung (R4), die mit dem Netzwerk gekoppelt ist, und so ausgebildet ist, daß sie einen ersten Datenbank-Identifizierer aus dem Inhalt einer ersten Datenbank erzeugt, welche der festgelegten Leitwegvorrichtung zugeordnet ist, wobei der erste Datenbank-Identifizierer eindeutig den Inhalt der ersten Datenbank repräsentiert; und zumindest eine andere Leitwegvorrichtung des Netzwerks, die so ausgebildet ist, daß sie einen zweiten Datenbank-Identifizierer aus dem Inhalt einer zweiten Datenbank erzeugt, welche der anderen Leitwegvorrichtung zugeordnet ist, wobei der zweite Datenbank-Identifizierer eindeutig den Inhalt der zweiten Datenbank repräsentiert, wobei die festgelegte Leitwegvorrichtung periodisch den ersten Datenbank-Identifizierer an die andere Leitwegvorrichtung verteilt, welche den ersten Datenbank-Identifizierer mit dem zweiten Datenbank-Identifizierer vergleicht, um zu bestimmen, ob die Datenbank-Identifizierer und daher die erste und zweite Datenbank konsistent sind.

6. Anordnung nach Anspruch 5, bei welcher die Datenbank-Identifizierer eindeutig vollständige Sequenznummernpakete (CSNP) (400) repräsentieren, welche den Inhalt von Verbindungszustandspaket- (LSP) (300) Datenbanken betreffen, und wobei die festgelegte Leitwegvorrichtung den ersten Datenbank-Identifizierer aus dem CSNP erzeugt, und periodisch den ersten Datenbank-Identifizierer, statt des CSNP selbst, an die andere Leitwegvorrichtung sendet.

7. Verfahren zum wirksamen Synchronisieren der Inhalte von Datenbanken, die auf Knoten (S1-S6;

R1-R6) eines Computernetzwerks (200) gespeichert sind, dadurch gekennzeichnet, daß das Verfahren umfaßt:

die Erzeugung eines ersten Datenbank-Identifizierers aus dem Inhalt einer ersten Datenbank, welche einem festgelegten Knoten zugeordnet ist, der mit dem Netzwerk gekoppelt ist, wobei der erste Datenbank-Identifizierer eindeutig den Inhalt der ersten Datenbank repräsentiert; 5

die Berechnung zweiter Datenbank-Identifizierer aus den Inhalten zweiter Datenbanken, die mehreren anderen Leitwegvorrichtungen zugeordnet sind, die mit dem Netzwerk gekoppelt sind, wobei die zweiten Datenbank-Identifizierer eindeutig den Inhalt der zweiten Datenbanken repräsentieren; 15

die periodische Sendung des ersten Datenbank-Identifizierers an die anderen Leitwegvorrichtungen über das Netzwerk; und 20

die Verarbeitung des ersten Datenbank-Identifizierers an den anderen Leitwegvorrichtungen, um zu bestimmen, ob irgendwelche Diskrepanzen zwischen der ersten Datenbank und den zweiten Datenbanken vorhanden sind. 25

8. Verfahren nach Anspruch 7, bei welchem der Schritt der Erzeugung den Schritt der Erzeugung des ersten Datenbank-Identifizierers umfaßt, welcher eindeutig ein vollständiges Sequenznummernpaket (CSNP) repräsentiert, welches den Inhalt einer Verbindungszustandspaket-(LSP) Datenbank betrifft, welche der festgelegten Leitwegvorrichtung zugeordnet ist. 30

9. Verfahren nach Anspruch 8, bei welchem der Schritt der Verarbeitung den Schritt des Vergleichs des zweiten Datenbank-Identifizierers mit dem ersten Datenbank-Identifizierer umfaßt, und dann, wenn der zweite Datenbank-Identifizierer zum ersten Datenbank-Identifizierer paßt, der erste Datenbank-Identifizierer an den anderen Leitwegvorrichtungen gespeichert wird, jedoch dann, wenn der zweite Datenbank-Identifizierer nicht zum ersten Datenbank-Identifizierer paßt, das CSNP von der festgelegten Leitwegvorrichtung angefordert wird. 35

10. Verfahren nach Anspruch 8, mit folgenden weiteren Schritten:

Erzeugung dritter Datenbank-Identifizierer (625a-c), die jeweils eindeutig ein Fragment (450a-c) des CSNP repräsentieren, welches den Inhalt einer LSP-Datenbank betrifft, die der festgelegten Leitwegvorrichtung zugeordnet ist; und 40

Berechnung vieter Datenbank-Identifizierer, die jeweils eindeutig ein Fragment des CSNP repräsentieren, welches den Inhalt einer LSP-Datenbank betrifft, welche jeder anderen Leitwegvorrichtung zugeordnet ist. 45

11. Verfahren nach Anspruch 10, bei welchem der Schritt der Verarbeitung den Schritt des Vergleichs des zweiten Datenbank-Identifizierers mit dem ersten Datenbank-Identifizierer umfaßt, und dann, wenn der zweite Datenbank-Identifizierer zum ersten Datenbank-Identifizierer paßt, der erste Datenbank-Identifizierer an den anderen Leitwegvorrichtungen gespeichert wird, jedoch dann, wenn der zweite Datenbank-Identifizierer nicht zum ersten Datenbank-Identifizierer paßt, ein bestimmter dritter Datenbank-Identifizierer mit einem entsprechenden vierten Datenbank-Identifizierer verglichen wird, um irgendwelche Diskrepanzen aufzulösen. 50

Revendications

1. Mécanisme pour synchroniser efficacement le contenu de bases de données stockées sur des noeuds d'un réseau d'ordinateurs (200), caractérisé en ce que le mécanisme comprend des moyens pour générer un identificateur de base de données (510) généré par chaque noeud du réseau d'ordinateurs à partir des contenus de bases de données associées à chaque noeud, l'identificateur de bases de données étant uniquement représentatif du contenu de la base de données associée à chaque noeud, d'où il résulte qu'un noeud (S1-S6 ; R1-R6) du réseau comprend des moyens pour distribuer son identificateur de base de données à un noeud de réception (D) du réseau et le noeud de réception comprend des moyens pour comparer son identificateur de base de données avec l'identificateur de base de données distribué pour déterminer si les identificateurs, et ainsi les bases de données, sont synchronisés. 55
2. Mécanisme selon la revendication 1, dans lequel l'identificateur de base de données est uniquement représentatif d'un paquet de nombres de séquences complet (CSNP) (400) appartenant au contenu d'une base de données de paquets d'états de liaison (LSP) (300) du noeud de distribution (S1-S6 ; R1-R6). 60
3. Mécanisme selon la revendication 2, dans lequel les noeuds sont des routeurs (R1-R6) et dans lequel le noeud de distribution est un routeur désigné 65

(R4).

4. Mécanisme selon la revendication 1, dans lequel l'identificateur de base de données (510) est généré par un algorithme d'un condensé de messages cryptographiques configuré pour transformer le contenu de la base de données en un condensé unique de longueur fixée, par un contrôle de redondance cyclique, ou par une génération de nombres de séquences. 5

5. Agencement pour synchroniser efficacement le contenu de bases de données stockées sur des routeurs (R1-R6) d'un réseau d'ordinateurs (200), caractérisé en ce que l'agencement comprend : 10

- un routeur désigné (R4) couplé au réseau et configuré pour générer un premier identificateur de base de données à partir du contenu d'une première base de données associée au routeur désigné, le premier identificateur de base de données étant uniquement représentatif du contenu de la première base de données ; et 15
- au moins un autre routeur du réseau configuré pour générer un deuxième identificateur de bases de données à partir du contenu d'une deuxième base de données associée à l'autre routeur, le deuxième identificateur de base de données étant uniquement représentatif du contenu de la deuxième base de données, 20

dans lequel le routeur désigné distribue périodiquement le premier identificateur de base de données à l'autre routeur, qui compare le premier identificateur de base de données au deuxième identificateur de base de données, pour déterminer si les identificateurs de bases de données, et ainsi les première et deuxième base de données sont consitants entre eux. 25

6. Agencement selon la revendication 5, dans lequel les identificateurs de base de données sont uniquement représentatifs de paquets de nombres de séquences complets (CSNP) (400) appartenant au contenu de bases de données de paquets d'états de liaison (LSP) (300), et dans lequel le routeur désigné génère le premier identificateur de base de données à partir du CSNP et diffuse périodiquement le premier identificateur de base de données, plutôt que le CSNP lui-même, à l'autre routeur. 30

7. Procédé pour synchroniser efficacement le contenu de bases de données stockées sur des noeuds (S1-S6 ; R1-R6) d'un réseau d'ordinateurs (200), caractérisé en ce que le procédé comprend les étapes suivantes : 35

- générer un premier identificateur de base de données à partir du contenu d'une première base de données associée à un noeud désigné couplé au réseau, le premier identificateur de base de données étant uniquement représentatif du contenu de la première base de données ; 40
- calculer des deuxièmes identificateurs de base de données à partir du contenu de deuxièmes bases de données associées à une multiplicité d'autres routeurs couplés au réseau, les deuxièmes identificateurs de base de données étant uniquement représentatifs du contenu des deuxièmes bases de données ; 45
- diffuser périodiquement le premier identificateur de base de données aux autres routeurs sur le réseau ; et
- traiter le premier identificateur de base de données au niveau des autres routeurs pour déterminer s'il existe des désaccords entre la première et la deuxième bases de données. 50

8. Procédé selon la revendication 7, dans lequel l'étape de génération comprend l'étape de génération du premier identificateur de base de données qui est uniquement représentatif d'un paquet de nombres de séquences complet (CSNP) appartenant au contenu d'une base de données de paquets d'états de liaison (LSP) associée au routeur désigné. 55

9. Procédé selon la revendication 8, dans lequel l'étape de traitement comprend l'étape consistant à comparer le deuxième identificateur de base de données avec le premier identificateur de base de données et, si le deuxième identificateur de base de données correspond au premier identificateur de base de données, le premier identificateur de base de données est stocké au niveau des autres routeurs, ou, si le deuxième identificateur de base de données ne correspond pas au premier identificateur de base de données, le CSNP est requis auprès du routeur désigné. 60

10. Procédé selon la revendication 8, comprenant en outre les étapes suivantes : 65

- générer des troisièmes identificateurs de base de données (625a-c) dont chacun est uniquement représentatif d'un fragment (450a-c) du CSNP appartenant au contenu d'une base de données de LSP associée au routeur désigné ; et
- calculer des quatrièmes identificateurs de base de données, dont chacun est uniquement représentatif d'un fragment du CSNP appartenant au contenu d'une base de données de LSP associée à chaque autre routeur. 70

11. Procédé selon la revendication 10, dans lequel 75

l'étape de traitement comprend l'étape de comparaison du deuxième identificateur de base de données avec le premier identificateur de base de données et, si le deuxième identificateur de base de données correspond au premier identificateur de base de données, le premier identificateur de base de données est stocké au niveau des autres routeurs, ou, si le deuxième identificateur de base de données ne correspond pas au premier identificateur de base de données, un troisième identificateur de base de données particulier est comparé à un quatrième identificateur de base de données correspondant pour résoudre tout désaccord.

5

10

15

20

25

30

35

40

45

50

55

11

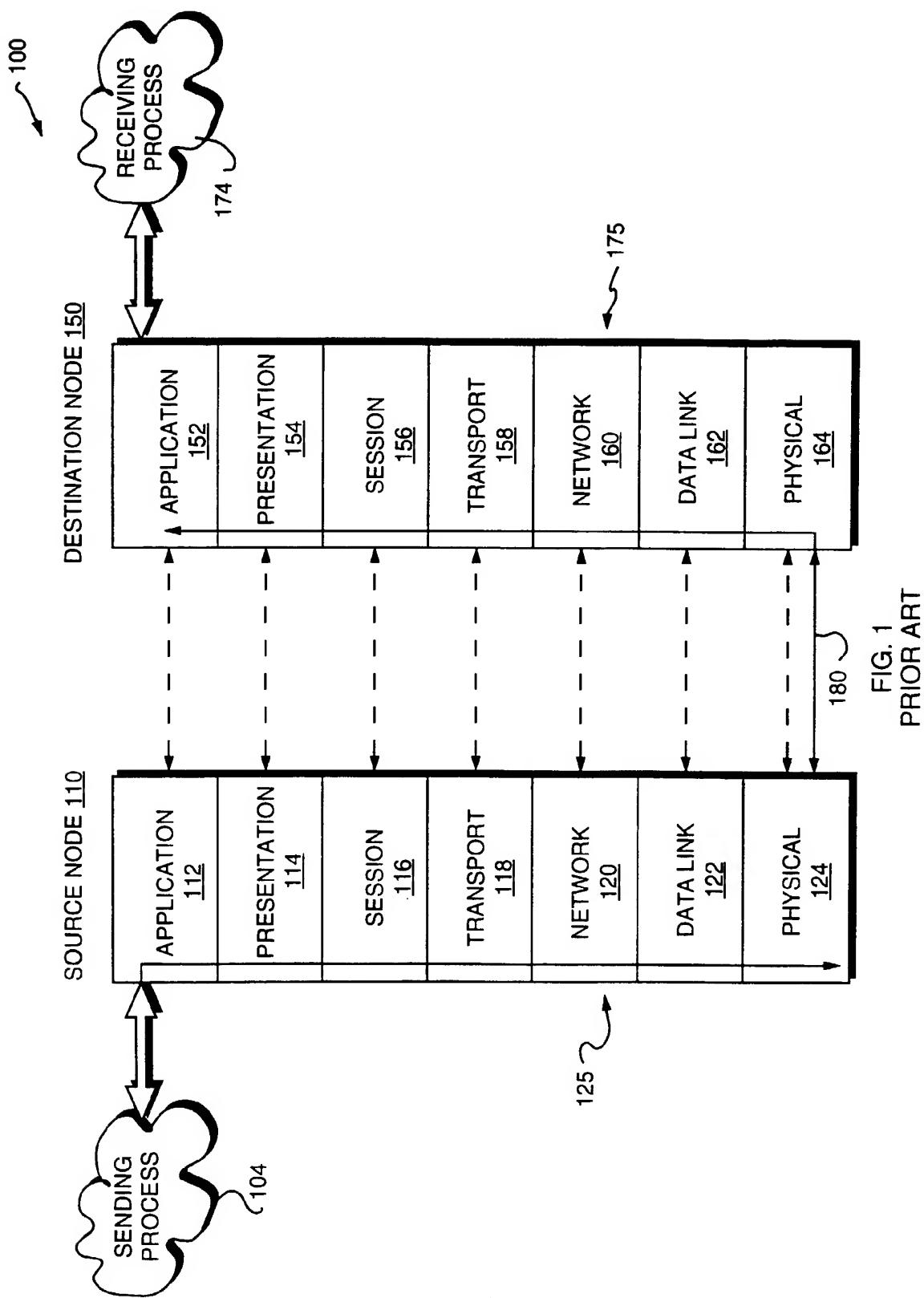
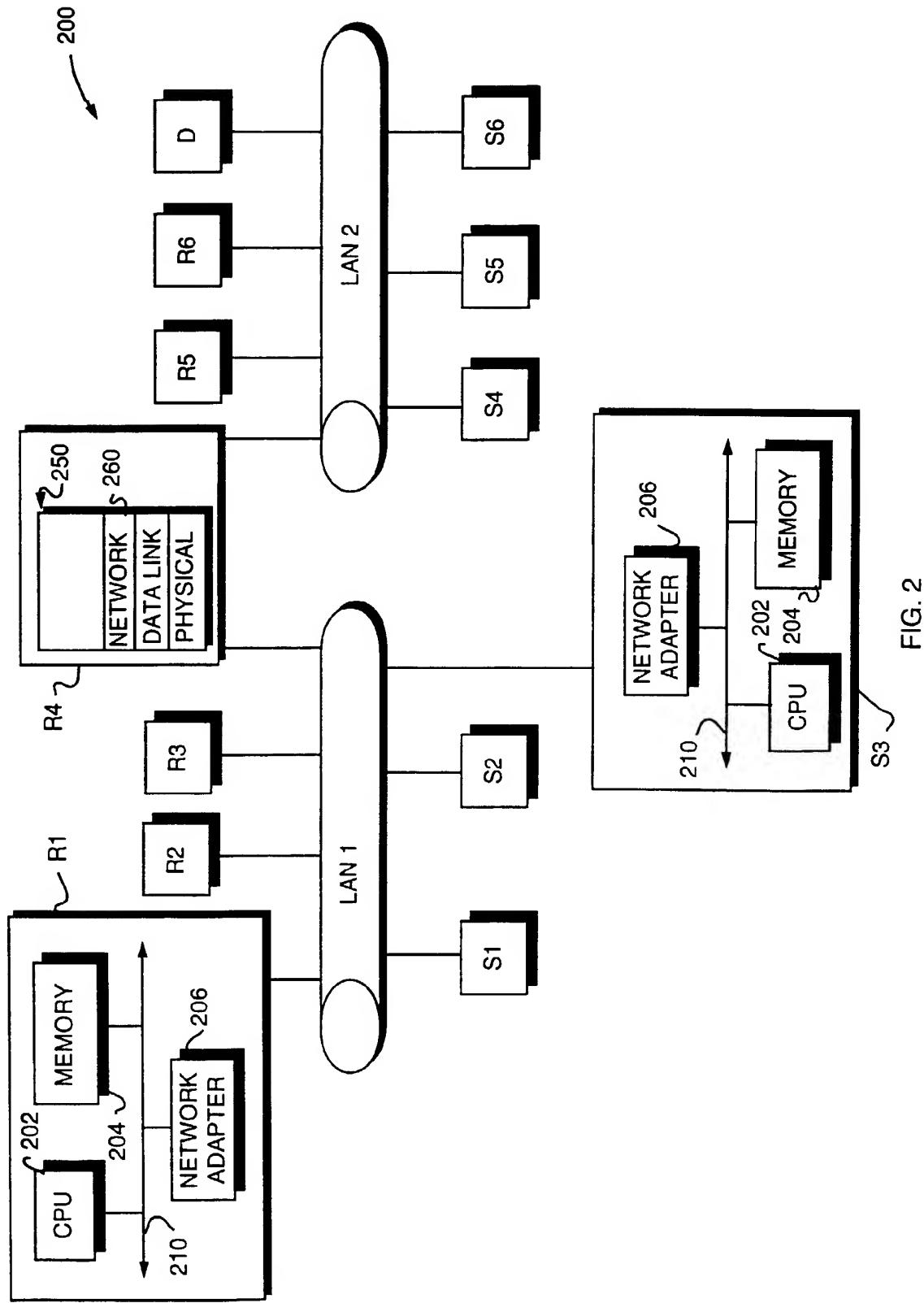


FIG. 1
PRIOR ART



300 ↗

| |
|----------------------------|
| SOURCE <u>302</u> |
| SEQUENCE NUMBER <u>304</u> |
| |
| NEIGHBORS <u>306</u> |

FIG. 3

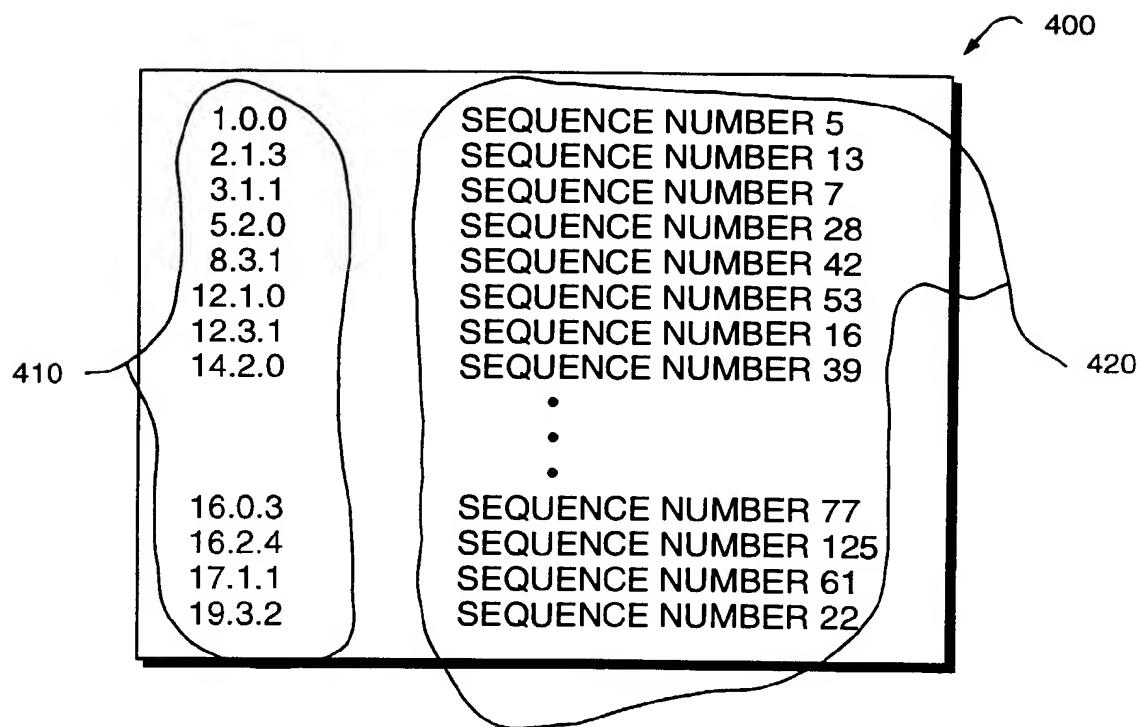
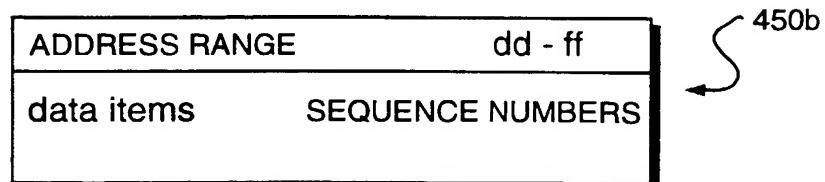
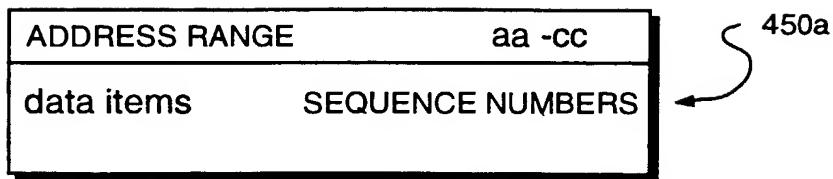


FIG. 4A



•

•

•

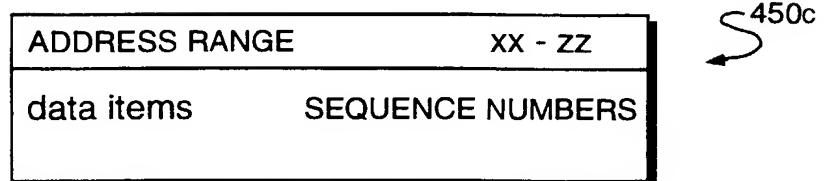
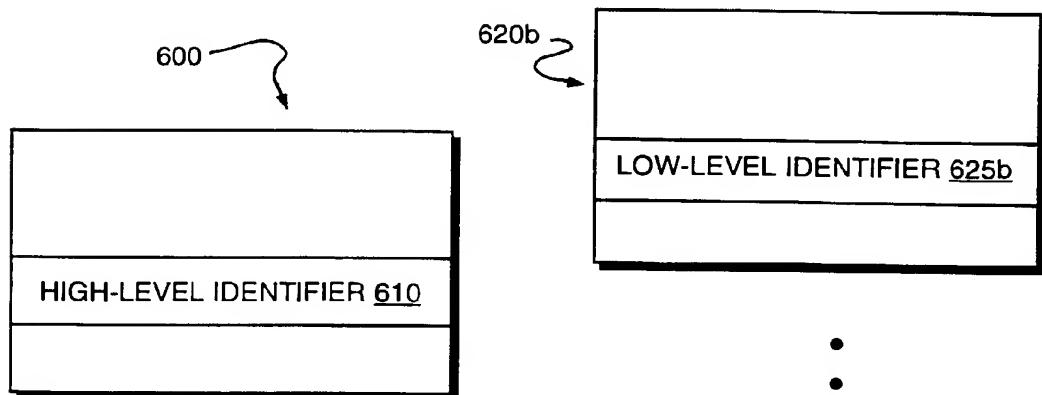
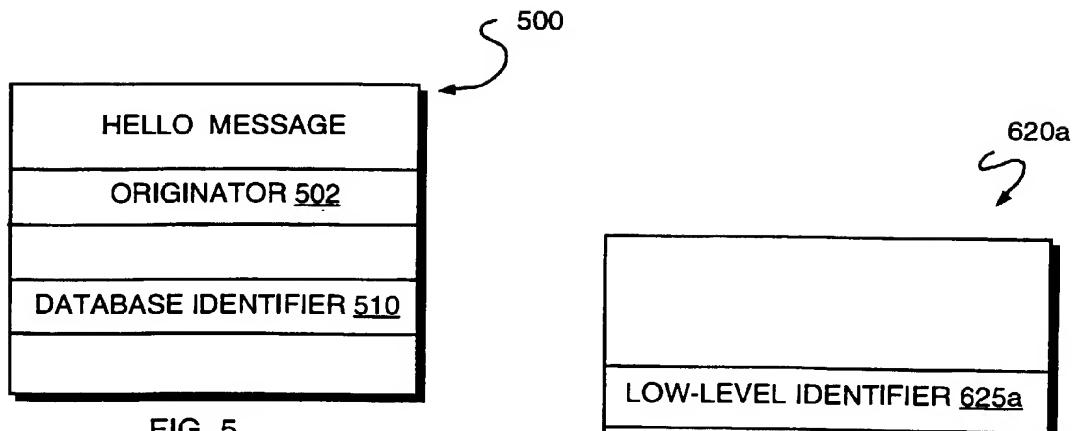
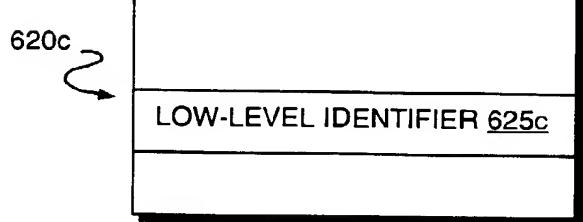


FIG. 4B



•
•
•



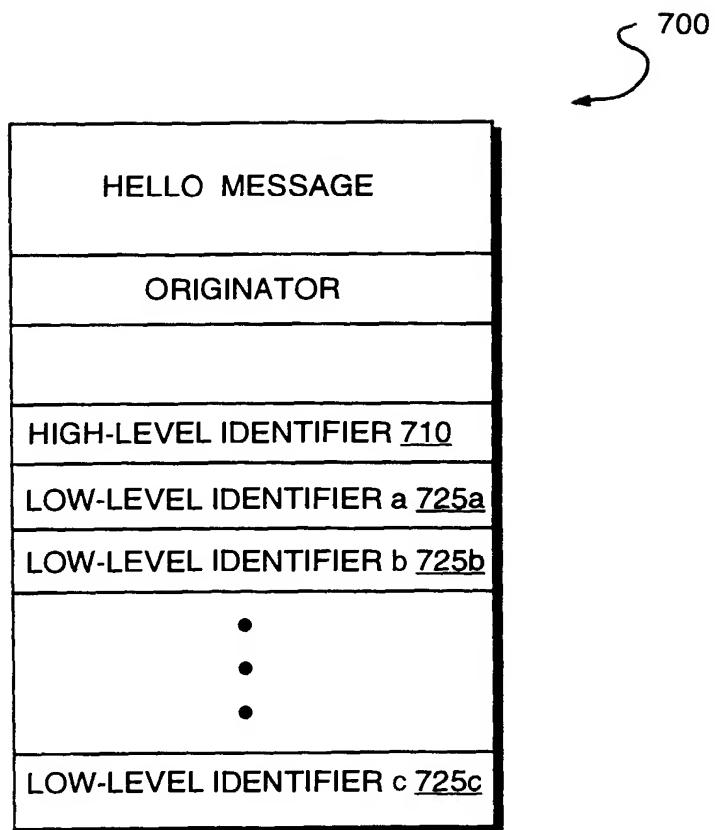


FIG. 7